Michel Fest
October 2010
1945-2010

UCSD

Stochastic Observability & Optimal Control

Marie-Antoinette Poubelle
Turramurra Institute of Technical Sciences

Bob Bitmead, UCSD

Brief Outline

Michel has a good story about underpants. You should ask him. I forget it. But I am certain it is a good story

Brief outline (for Michel)

Brief outline (for Michel)

- TCP network traffic congestion control
  - A gentle introduction
  - An observability question arises in stochastic (HMM) systems
- Linear systems definitions; deterministic and stochastic
  - Flexing muscles
  - Covariance based ideas
- Nonlinear system definition
  - Conditional entropy ideas replace covariances
  - Careful consideration of *complete* reconstructibility
- Return to TCP congestion control
- Reconstructibility and optimal control
- Connections to Michel's work in identification and control

A heads-up

There are lots of control problems in communications at the many different system layers; power control, admission control, congestion control

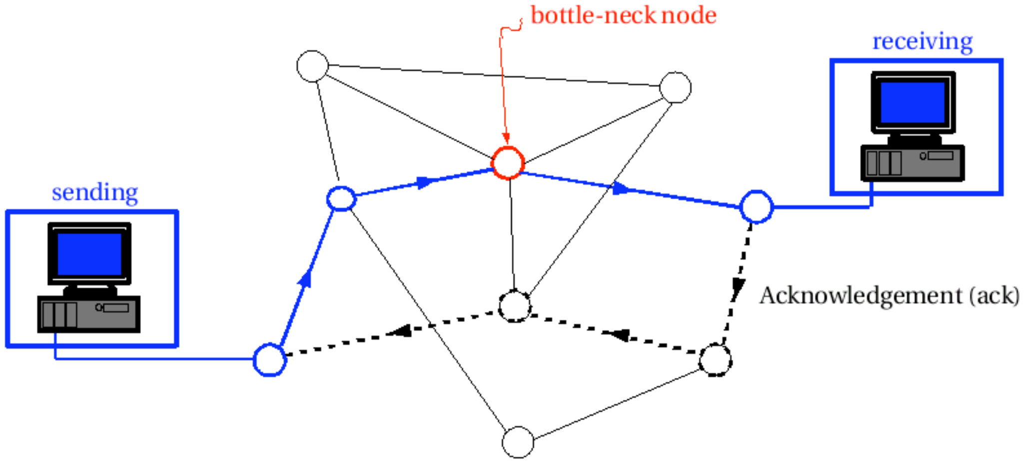
We wish to look at network congestion control in a stochastic environment

A natural state estimation problem arises

What do we mean by stochastic observability?

This is our story ...

TCP as feedback



From the perspective of the source node

packets are sent into the network

a packet arriving at destination produce an ACK packet response

the send rate is controlled into network based on ACK sequence

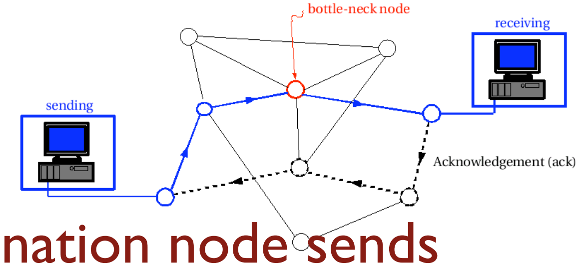
the aim is to avoid traffic congestion

the competing traffic is stochastic in nature

normally modeled as dominated by a single bottleneck node

This is a stochastic feedback control problem

Destination node behavior



Upon receipt of a packet from the source, the destination node sends an acknowledgement packet in reply

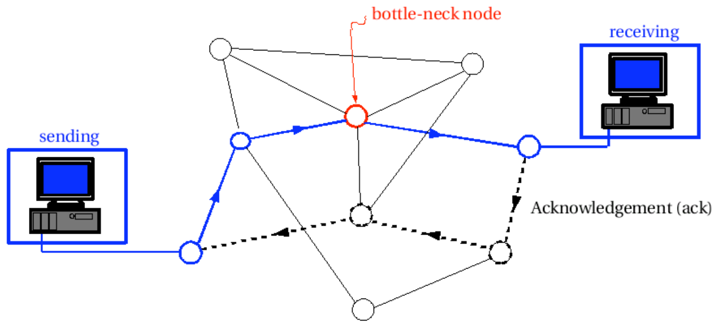
- Generally ACKs are simple few-bit packets
  - arriving packet identifier

- There are proposals for ACKs to contain more and more useful data
  - arrival time
  - data inserted by intervening nodes
    - buffer state and/or statistics
    - traffic state and/or statistics

The ACKs have to travel back through the same network

- They too can be lost

Source node behavior

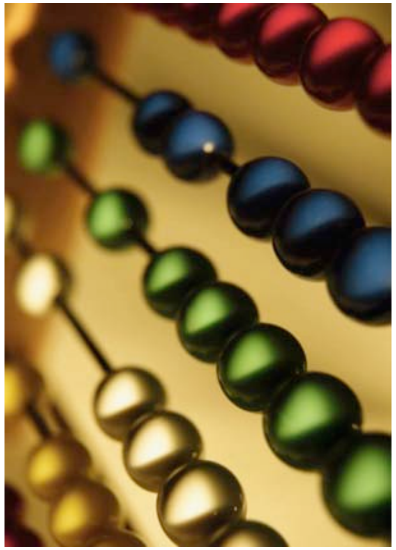


Data made up into packets with rate r_k

Packet rate into network is the mechanism of congestion control

Rate r_k is adjusted in response to arrival or non-arrival of ACKs

non-arrival: time-out or ACK out of sequence



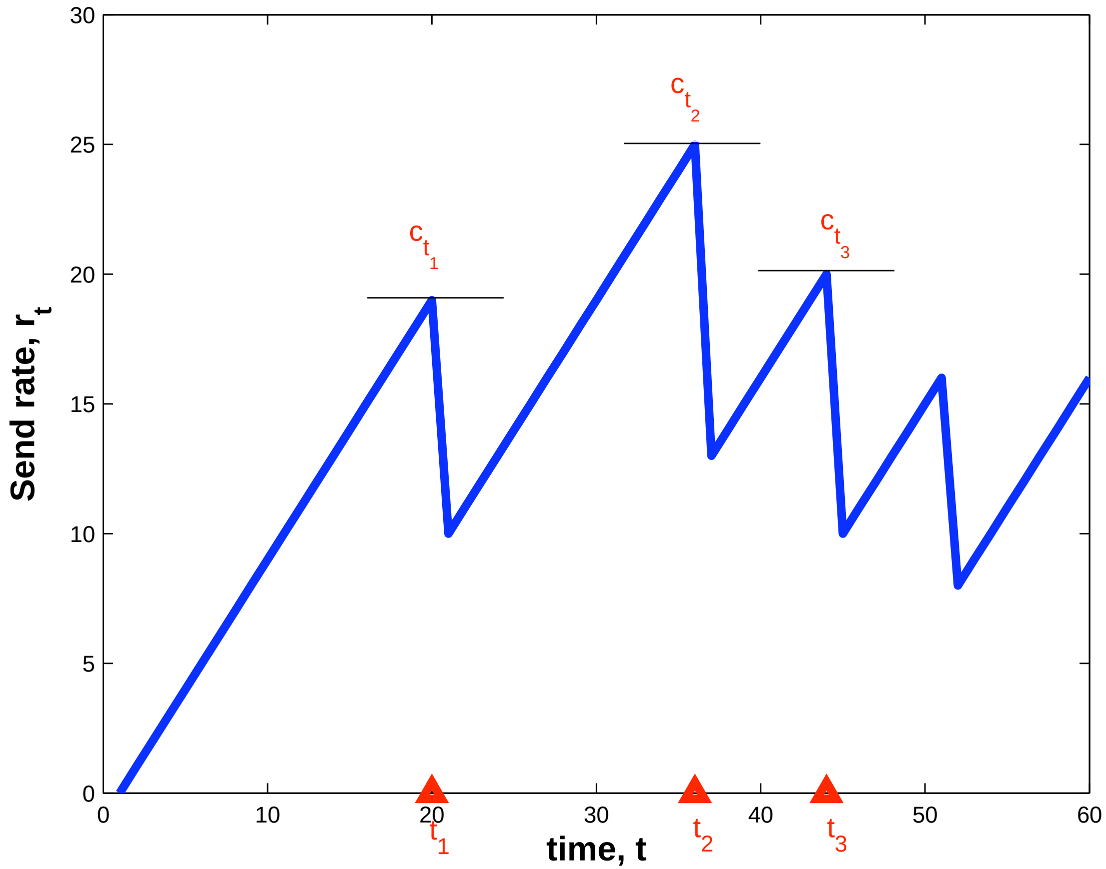
Common congestion control law AIMD

Additive increase / multiplicative decrease

ACK arrives: increase rate r_k by one packet

ACK missing: decrease r_k by a factor of two

AIMD

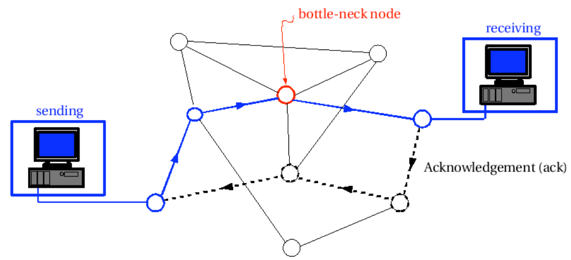


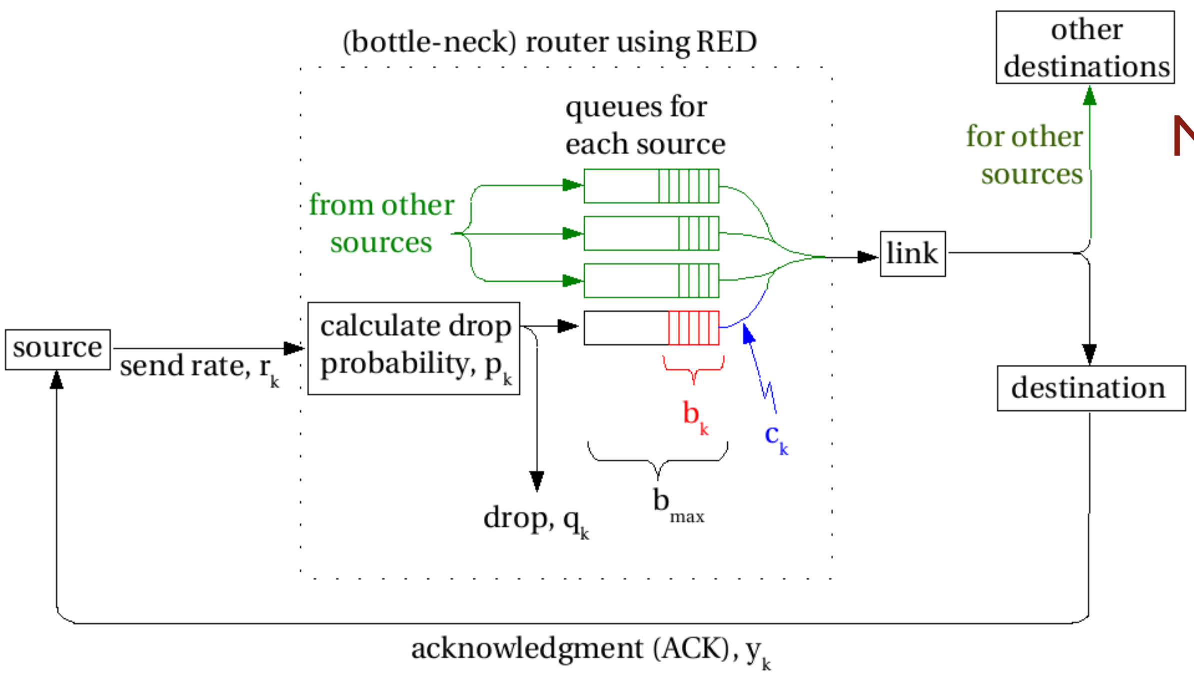
Additive increase / multiplicative decrease

ACK arrives: increase rate r_k by one packet

ACK missing: decrease r_k by a factor of two

Bottleneck node behavior





Nodes have finite buffers

- buffer occupancy b_k
- competing traffic (random) c_k
- arrival rate from source r_k
- packets deliberately dropped q_k

Packets are dropped with a certain probability p_k

Random Early Detection (RED) algorithm

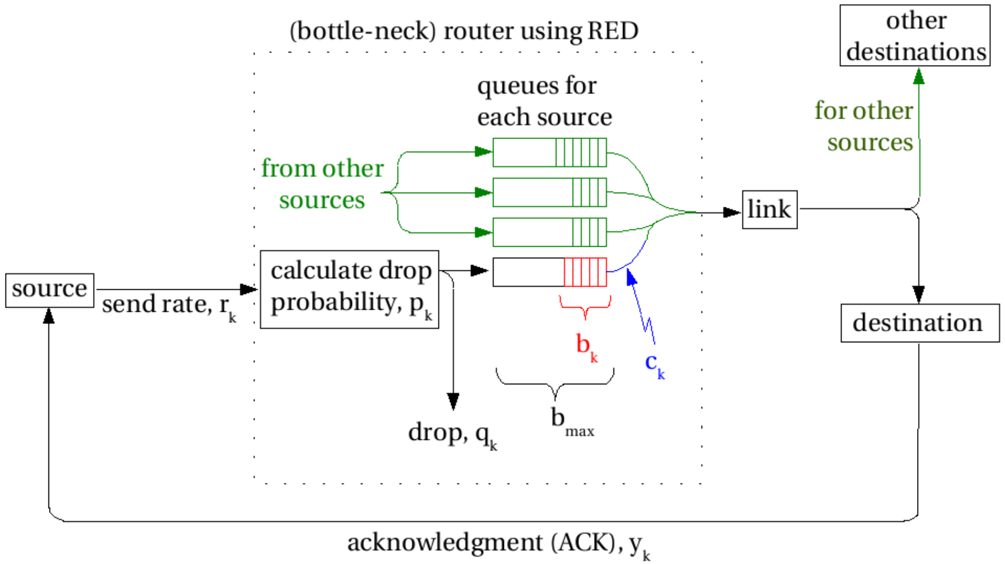
p_k depends on buffer state b_k

$$\begin{array}{lcl} b_k \to \text{full} & \Longrightarrow & p_k \to 1 \\ b_k \to \text{empty} & \Longrightarrow & p_k \to 0 \end{array}$$

Droptail algorithm drops all packets when buffer full

There are other algorithms

Hidden Markov Model of Bottleneck Node



State of the bottleneck node

$$x_k = \begin{pmatrix} b_k \\ c_k \end{pmatrix}$$

Drop probability $\quad p_k = f(b_k)$

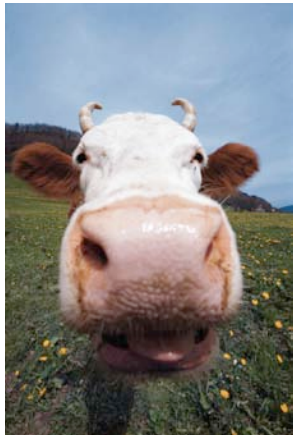
Capacity (traffic) model - Markov chain

$$P(c_{k+1} = c) = \sum_d P(c_{k+1} = c | c_k = d) P(c_k = d)$$

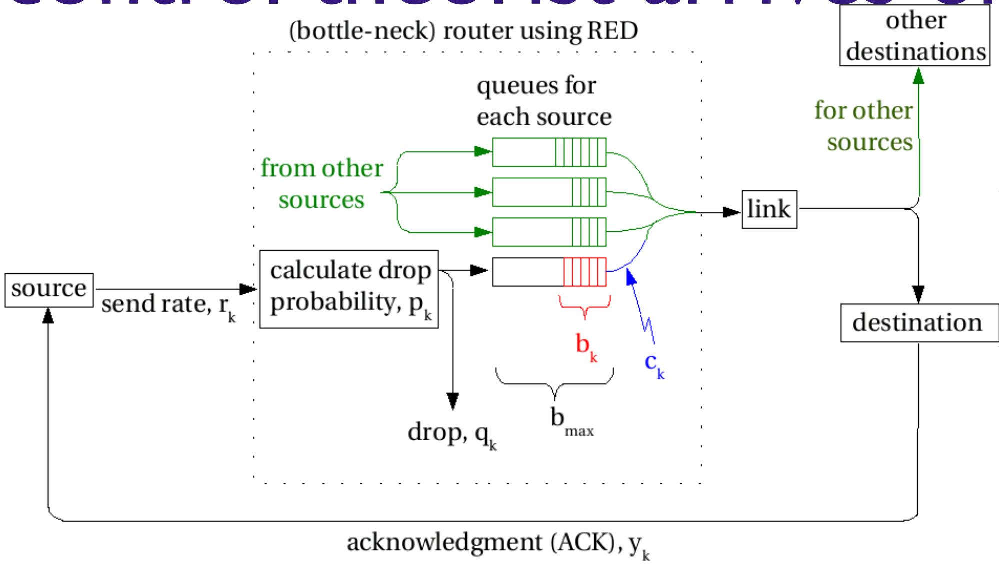
Hidden Markov Model for bottleneck node state

$$\begin{aligned} \Pi_x(k+1) &= A(r_k)\Pi_x(k) \\ \Pi_y(k) &= C\Pi_x(k) \end{aligned}$$

Known model - input sequence r_k and output (ACK) sequence y_k



A control theorist arrives on the scene



- Can the source computer reliably estimate the state of the bottleneck node?
  - Buffer length and capacity value
    - Available data are input rate history and ACK sequence history
- This is an observability question about an HMM
  - A theory of HMM filtering and smoothing exists
    - But what about the observability questions?

Does the input-output sequence suffice to estimate the state (c_k, b_k)?

Stochastic observability

Deterministic linear system

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k \end{aligned}$$

Complete observability: given measurements $\{u_0, y_0, \ldots, u_{n-1}, y_{n-1}\}$
we can compute the initial state, x_0, exactly

Complete reconstructibility: given measurements $\{u_0, y_0, \ldots, u_{n-1}, y_{n-1}\}$
we can compute the current state, x_{n-1}, exactly

Stochastic linear system

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + v_k \\ y_k &= Cx_k + w_k \end{aligned}$$

What do we now mean?

Generally, there is no exact answer

Suggested (gaussian) stochastic linear system definition

Assume the initial state is gaussian $\mathcal{N}\left(\bar{x}_0, \Sigma_{0|-1}\right)$. If, for any vector ξ
either: $\xi^T \Sigma_{0|-1} \xi = 0$ or $\xi^T \Sigma_{0|n-1} \xi < \xi^T \Sigma_{0|-1} \xi$
then the system is completely observable

Stochastic reconstructibility

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + v_k \\ y_k &= Cx_k + w_k \end{aligned}$$

Assume the initial state is gaussian $\mathcal{N}\left(\bar{x}_0, \Sigma_{0|-1}\right)$. If, for any vector ξ

either: $\xi^T \Sigma_{n-1|-1} \xi = 0$ or $\xi^T \Sigma_{n-1|n-1} \xi < \xi^T \Sigma_{n-1|-1} \xi$

then the system is completely reconstructible

Property: the system above is completely observable iff is full rank

$$\mathcal{O} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix}$$

Property: the system above is completely reconstructible if $\mathrm{Range}(\mathcal{O}_n) \supseteq \mathrm{Range}(A^n)$

Gaussian Systems - proving things

$$\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix} x_0 + \begin{pmatrix} 0 & 0 & \dots & 0 \\ CB & 0 & \dots & 0 \\ CAB & CB & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{n-1}B & CA^{n-3}B & \dots & CB \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_{n-1} \end{pmatrix} + \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ \vdots \\ v_{n-1} \end{pmatrix}$$

Measurement vector and initial state are jointly gaussian

Simple formula for conditional mean and conditional variance for x_0

$$\Sigma_{0|n-1} = \Sigma_0 - \Sigma_0 \mathcal{O}^T \left(\mathcal{O} \Sigma_0 \mathcal{O}^T + \mathcal{H} \mathcal{Q} \mathcal{H}^T + \mathcal{R} \right)^{-1} \mathcal{O} \Sigma_0$$

Strict inequality requires

$$\Sigma_{0|-1} > 0 \qquad \operatorname{rank}(\mathcal{O}) = n \qquad R < \infty \qquad Q < \infty \quad (n > 1)$$

These ideas extend easily to non-gaussian systems

Extension to Nonlinear Stochastic Systems

$$y = |x|$$

Is x observable from y?

Depends on the distribution of x

If the distribution is symmetric then $\mathrm{sgn}(x)$ is unobservable

Example 1: $x = \begin{cases} -1 & \text{, with probability } 1/2 \\ 1 & \text{, with probability } 1/2 \end{cases}$ y unobservable

Example 2: $x \sim \mathcal{N}(1, \sigma^2)$ y observable

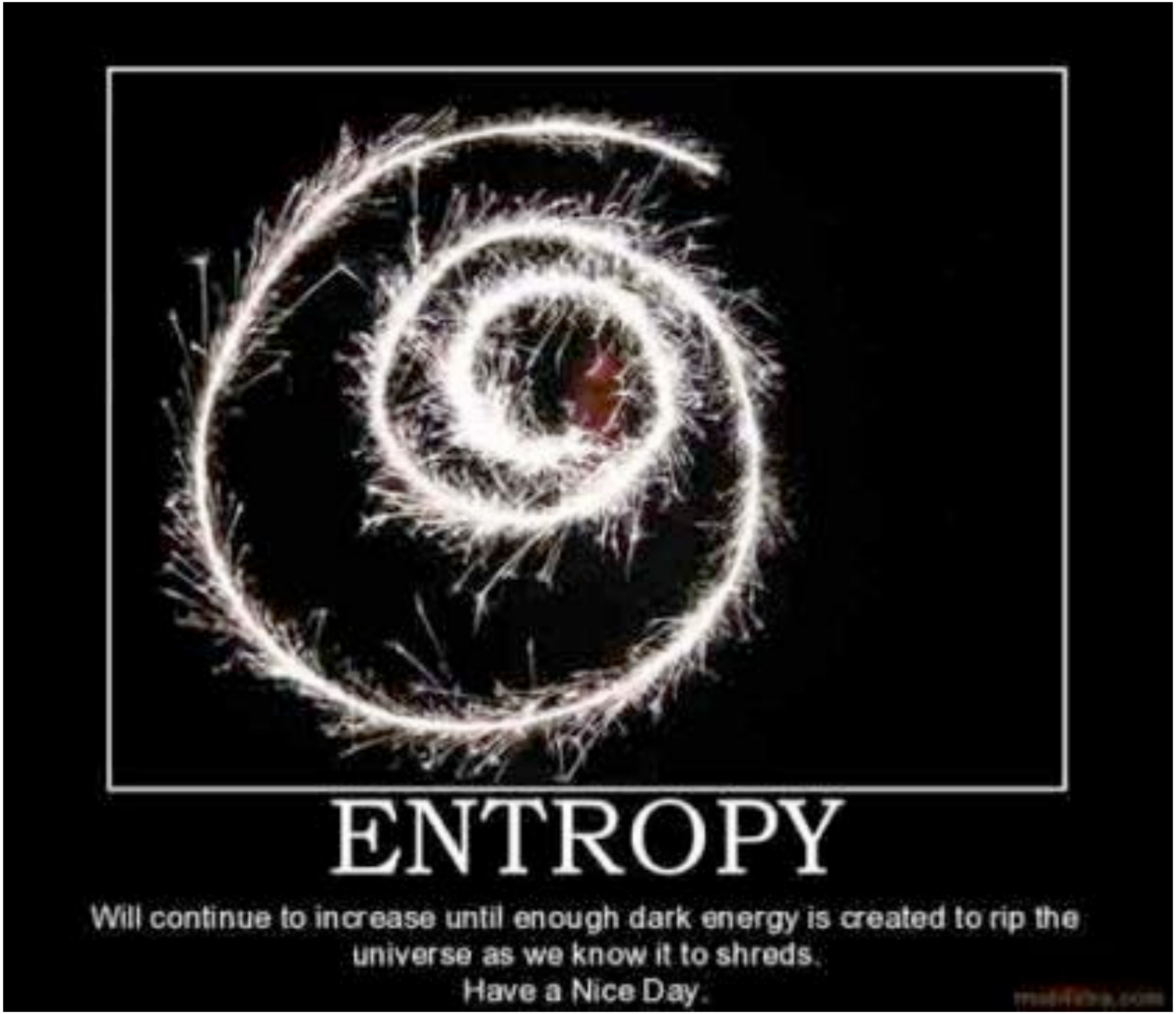
Ideas of complete observability need rethinking

There is a need to consider the observability of functions of x

Variance might not be the correct quantity to consider

Especially for HMMs where we are reconstructing an estimate of the probability density

Use entropy $H(x)$ and conditional entropy $H(x|y)$


ENTROPY
Will continue to increase until enough dark energy is created to rip the
universe as we know it to shreds.
Have a Nice Day.

Conditional Entropy

$$\begin{aligned} H(x|y) &= \sum_{y_i\in\mathcal{Y}} P(y=y_i) \sum_{x_j\in\mathcal{X}} P(x=x_j|y=y_i) \ln P(x=x_j|y=y_i) \\ &= H(y,x) - H(x) \end{aligned}$$

Properties of conditional entropy

$H(x) \geq 0$

$H(g(x)) \leq H(x)$ with equality if $g(.)$ is injective

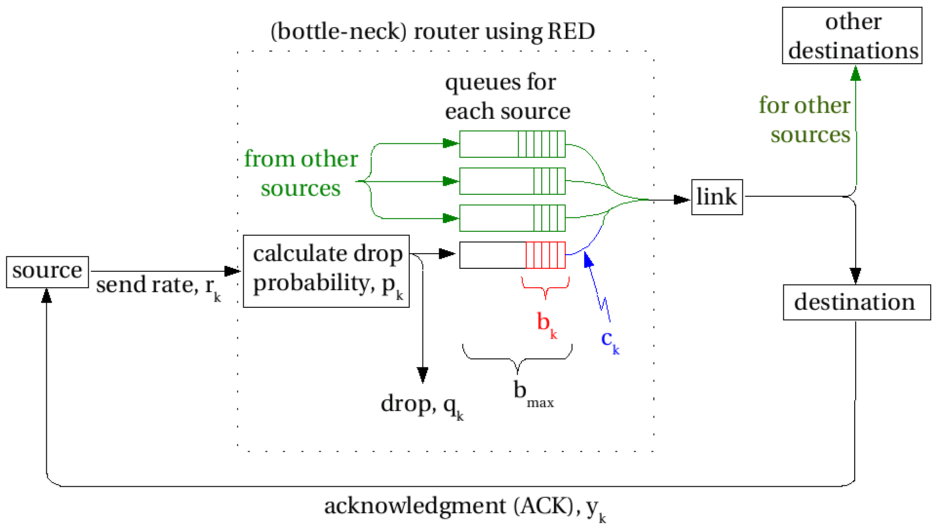
$H(x|y) \leq H(x)$ with equality iff x and y are independent

$H(x) = \mathrm{tr}(\Sigma)$ if x is gaussian $N(\bar{x}, \Sigma)$

Definition: Random quantity x is completely observable from random quantity y if, for every, measurable function $g(.)$:$\mathcal{X} \rightarrow I\!R$, either $H(g(x))$=0 or $H(g(x)|y)<H(g(x))$

Testable approach for finite state systems like HMM

Back to the internet



HMM observability

$$
\begin{aligned}
\Pi_x(k+1) &= A(r_k)\Pi_x(k) \\
\Pi_y(k) &= C\Pi_x(k)
\end{aligned}
$$

We need to test the conditional entropy of every function of the HMM state

- Since the output (ACK) sequence is denumerable and the states are denumerable, the number of tested functions is finite
  - The tests can be exhaustively evaluated
    - The test protocol depends on the specific control (source rate) sequence in operation

The bottleneck node state is observable/reconstructible from the source when operating with the TCP/IP control

Reconstructibility and optimal control

Reconstructibility and optimal control

For Hidden Markov Model systems

Theorem: if the system is unreconstructible no matter what control is applied then the closed-loop optimal control law is the same as open-loop optimal control law

Theorem: if the conditional entropy of the reward function given the inputs and outputs equals that given the inputs alone then the closed-loop optimal control law is the same as the open-loop

So what?

In order to achieve optimal feedback control the control law needs to be able to expose the state or at least the reward function value

This is the familiar question of Dual Adaptive Control

A tension between excitation versus regulation

Connection to Michel Gevers' works

Michel has studied connections between system identification & control

- Mostly in stochastic contexts

Identifiability is a subset of reconstructibility (parameters as states)

- Michel: experiment design for identification for control
  - Management of model error quality
- Michel: iterative identification and control design
  - Successive optimization
- Michel: iterative feedback tuning
  - Excitation issues

Results presented here attack but do not (yet) solve the underlying problem

Too bad!

Bozos

Bonne retraîte Michel!!!

de tes grandes amis yanquis

et de ta plus grande amie incarcérée

Michel, please explain

Michel, please explain

FIN